# Vote and Voice: An experiment on the effects of inclusive governance rules

Shaun P. Hargreaves Heap (King's College, London)

Kei Tsutsui (University of Bath)

Daniel J. Zizzo (University of Queensland)*

**Abstract**

We present an experiment that examines three mechanisms through which the extent of inclusivity in an organization's governance arrangements might affect its performance. We distinguish extent of inclusivity along two dimensions: members of the organization may or may not be able to a) vote on collective decisions ('vote') and b) discuss with others what should be done ('voice'). We find that the inclusivity can affect performance and that each dimension of inclusivity matters, but for different decision problems within an organization. The 'voice' matters for motivation whereas 'voting' matters for processing and aggregating information; and the decisive difference for performance comes from 'voice', not 'voting'.

*Keywords*: democracy, dictatorship, rights, cooperation, wisdom of crowd, rational ignorance.

*JEL Classification Codes*: C72, C91, C92, D43, H21.

*Corresponding author

## 1. Introduction

One consideration in the design of any organization or society is whether to involve its members in collective decision making.[1] Members can be included in two distinct ways. First, a person's view may or may not be formally incorporated in the collective decision making process: in effect, do they get to 'vote' on collective decisions? This dimension of inclusivity may range in a society from one person in a pure dictatorship or hierarchy to every adult in a democracy. Likewise, a collective decision in an organization might be made by an individual manager or by a committee. The other dimension of inclusivity relates to whether a member can exercise 'voice'. Do they have the opportunity to express an opinion in a discussion over the collective decision? These opportunities are determined, for example, in society by its regime of rights, notably with respect to free speech. In organizations, they are determined by the extent of internal practices of formal and informal consultation. In this paper we examine with a laboratory experiment how the dimensions of 'vote' and 'voice' inclusivity in governance rules affect the performance of an organization.

There is a large related non-experimental literature. For example, there are the literatures on whether democracy promotes development (e.g. see Przeworski and Limongi 1993; Helliwell 1994; Acemoglu et al 2008) and whether giving more voice in the Hirschman (1970) sense to employees improves organizational performance (e.g. see Ichniowski and Shaw 2003 and Black and Lynch 2004). This is an important body of work but three problems almost always surface in relation to the interpretation of the evidence on inclusivity and performance in such studies. Our choice of a laboratory experiment is a direct response to these problems.

[1] Of course, there is also a design choice over whether to make collective decisions: that is, there is always the alternative of forsaking collective decisions by devolving the domain of decision to individual decision makers, as in markets.

The first problem is the difficulty in establishing causality in what is often a simultaneous relation (e.g., between democracy and growth).[2] Experiments in the laboratory are well suited to tackling this because they offer control over what changes.

Second, successful inclusive organizations or societies, in practice, often bundle together the 'voting' and 'voice' aspects of inclusivity. As a result, the possible distinctive contributions of each dimension of inclusivity are often difficult to disentangle (e.g., see Mukand and Rodrik, 2015, who argue that rights - ie 'voice', and not 'votes' - play the crucial role). We tackle this by having a 2x2 design. Along the 'vote' dimension, either everyone votes as in a democracy (D) or one person within the organization votes (i.e. decides) on the collective action, as in a purely hierarchical organization (H) run by a dictator/manager. Along the 'voice' dimension, either everyone can express an opinion (Chat) or there are no discussion opportunities. Thus, we consider 4 polar types of governance in our treatments: Democratic (D) and Hierarchical (H) in 'voting' and each can be accompanied by 'voice' (=D+Chat and H+Chat, respectively) or not (= D and H). This design enables us to disentangle the distinct contributions of 'voting' and 'voice' in the performance of an organization.

The third problem we are able to address through the use of experiments is the mechanisms connecting inclusiveness with performance. In the non-experimental literature the relation between inclusiveness and performance often appears like a black box, with the precise mechanisms and their relative importance seemingly unclear. It is through the selection of decision problems in the experiment that we can test for precise possible mechanisms linking governance rules with organizational performance. A change in the 'vote' and 'voice' governance rules could affect performance through three distinct and independent mechanisms, that our experimental design will allow us to test separately. The mechanisms relate to

[2] See Glaeser et al (2004) for an argument doubting that the quality of institutions exercise an independent effect on performance.

*motivation, incentives and composition* within an organization. The three hypotheses, H1, H2 and H3, address respectively each of these mechanisms and come in two versions that pick up on the two governance dimensions: for example H1a addresses how the 'voting' dimension of governance affects performance through motivation and H1b does the same for the 'voice' dimension of governance. For each mechanism, we appeal to a central argument in the wider literature on how through this mechanism 'vote' and 'voice' might affect performance in an organization to generate the specific hypotheses.

The first possible mechanism through which inclusivity might affect performance is motivational. In particular, when discussing inclusivity in society, in *Considerations on Representative Government* (1861) John Stuart Mill argues that combining liberal freedoms with the extension of the franchise in representative government encourages people to take account of general interests and not just their own.

> 'The maximum of the invigorating effect of freedom upon the character is only obtained when the person acted on either is, or is looking forward to becoming, a citizen as fully privileged as any other.....Still more salutary is the moral part of the instruction afforded by the participation of the private citizen, if even rarely, in public functions. He is called upon, while so engaged, to weigh interests not his own; to be guided, in case of conflicting claims, by another rule than his private partialities; to apply, at every turn, principles and maxims which have for their reason of existence the general good.....' (Mill 1861, p.66/7)

Mill was echoing a very similar claim by De Tocqueville (1840) in *Democracy in America*. For convenience, we associate them with the 'voting' dimension of the motivational mechanism, even though their arguments have elements of 'voice'. We associate the 'voice' motivational argument with Hirschman (1970). He specifically identifies 'voice' as the source of a motivational change linked to 'loyalty' to an organization, where 'loyalty' involves a

person weighing the interests of the organization as a whole as well as their own when acting individually. He is not alone in making this argument (e.g., see Pateman 1970, Bowles and Gintis 1987, 1993; Dryzek and List, 2003); and there is some evidence on attitudes that is consistent with the claim (e.g., see Long 1978 and Frey 1998). To test these arguments, we examine whether the extension of inclusiveness via 'voting' and 'voice' in collective decision encourages cooperation in a *separate* public goods games. We focus on a decision arena that is different from the one of collective decision making because this is the claim of both Mill and DeTocqueville: that is, there are spillovers from participation in collective decision making into other arenas.[3] We use a public goods game as this other non-collective decision making arena because many interactions within an organization (e.g. over individual efforts which cannot be easily monitored) take the form of a public goods game and the efficiency of the organization will depend in part on the extent of cooperation in these interactions. H1a and H1b follow in relation to this motivational mechanism.

*H1a (Mill/deTocqueville) 'Voting' governance inclusiveness encourages cooperation in public goods games within an organization.*

*H1b (Hirschman) 'Voice' governance inclusiveness encourages cooperation in public goods games within an organization.*

The second inclusivity mechanism revolves around incentives. When the rules governing how collective decision are made change, this can affect the incentives individuals face to acquire information and skills. Downs (1957), in particular, makes the famous argument that the voting aspect of democracy discourages information acquisition.

'The marginal return on information acquired for voting purposes is measured by the expected gain from voting "correctly" instead of "incorrectly". However unless his vote

[3] The quote from Mill makes this spillover clear because the putative effect is upon a person's moral character. DeTocqueville similarly argues that it is the experience of participation in collective (political) decision making that spawns cooperation in the civic associations formed by individual decision making.

> actually decides the election, it does not cause the "right" party to be elected instead of the "wrong" party…Therefore, voting "correctly" produces no gain in utility whatsoever…Therefore we reach the startling conclusion that it is irrational for most citizens to acquire political information for the purposes of voting' (p.146-7).

In our experiment, subjects do not acquire information but they do have to process information that is made available to them and Downs's argument applies equally to the processing of information. Dictator/Managers are, in comparison, decisive and so have a stronger incentive to process information. H2a follows.

The influence of 'voice' on processing information is likely to encourage effort in processing information. This is because 'voice' generates a new source of incentive to process information: the scope to persuade others to their view with 'voice'. This is a new way in which a person can become decisive in collective decisions. H2b follows.

*H2a (Downs) 'Voting' governance inclusiveness discourages effort in processing information in collective decisions for the organization.*

*H2b 'Voice' governance inclusiveness encourages effort in processing information in collective decisions for the organization.*

The third mechanism through which inclusivity might affect organization performance is compositional. Different governance arrangements, by involving different people in different ways, affect the skill, information and interests that enter into a collective decision. In particular, Francis Galton, in effect and echoing Condorcet's Jury Theorem, famously argues, in his Nature essay, '*Vox Populi*', that inclusivity along a voting dimension helps a group to answer an informational problem better than can a typical individual: i.e. guessing the weight of a slaughtered and "dressed" ox. He focuses on the median guess (because of its connection to outcomes under democratic voting rules) and finds that it is surprisingly close to the actual weight.

> 'It appears then, in this particular instance, that the *vox populi* is correct to within 1%...and that individual estimates are distributed… in such a way that it is an equal chance whether one of them selected at random falls within or without the limits of -3.7% and +2.4%.... This result is, I think, more creditable to the trustworthiness of a democratic judgment than might have been expected' (Galton 1907, p.451).

This insight has become known as the wisdom of crowds and has recently been popularized by Surowiecki (2004) and Page (2007). It turns on a mathematical property of aggregating *independent* and diverse individually held information through 'voting'. H3a follows.

The arguments in the literature regarding the influence of 'voice' on information aggregation are different and contested. On the one hand, at least since Milton, there is the argument that the ability to put ideas into competition with each other that comes with the extension of 'voice' has beneficial effects for truth: 'whoever knew Truth put to the worse, in a free and open encounter?'. But against this, there is the worry that when people come together and discuss an issue a process of 'group-think' may take opinions in a disastrous direction (e.g., see Le Bon, 1895/2009). H3b takes Milton's side because it is a founding belief in liberal democracies that the exchange of ideas that comes with free speech is more likely to have these beneficial effects than the adverse one.

*H3a (Galton) 'Voting' governance inclusiveness improves the quality of collective decisions in an organization.*

*H3b (Milton) 'Voice' governance inclusiveness improves the quality of collective decisions in an organization.*

To test these hypotheses, we have a design where there are two groups of individuals in the laboratory who make two decisions under our four possible governance arrangements. One is an individual decision over contributions to an own group public good that affects the costs

of production for that group. We use this to test H1, the motivational mechanism. The other decision is a collective decision for each group. It is the price set by that group in what is a competition with the other group. This collective price decision affects the revenue coming to the group in which all group members share equally. In Democratic groups everyone makes a price suggestion and the median is selected. In Hierarchical groups, a single person (dictator/manager) makes the price suggestion and this is decisive; and, to complete the governance mix, the price decision can be accompanied by 'voice' or 'no voice'. The quality of the price suggestions and the price in this first decision can be judged by their distance from the best price response. We use this measure of the quality of individual price suggestions under the different governance arrangements to test H2 regarding the influence of incentives. Taking account of any differences in the quality of price suggestions revealed in testing H2 across governance arrangements, we can then use the quality of the actual group price decisions to test H3 for the influence of the composition mechanism under different governance arrangements.

The individual public goods decision is independent of the price setting decision. This is crucial for the separate identification of the motivational mechanism as contrasted with the incentive and compositional ones. Nevertheless, while the individual public goods decision is independent of the price setting decision, individuals are affected by the price decision because both public good contribution and price suggestions are (at least ex ante) expected to affect individual pay-offs: that is, they are incentivized. The disadvantage of the design is that the natural frame for our groups is then one of a specific kind of profit seeking organization. There is also an advantage of having a contextual frame, however, in that it may facilitate understanding of the instructions (Alekseev et al., 2017).

We are not the first to use experiments to examine how 'voting' and 'voice' affect collective decision making (e.g. Sutter, 2009; Bornstein et al, 2004; Brandts, 2014; Frohlich et

al., 1998). But to our knowledge, with one exception (Ellman and Pezanis-Christou, 2010, discussed in section 2), the comparison in these experiments is always either with the alternative exogenous determination of the collective decision (i.e., its determination by nature) or with what happens when there is no machinery for collective decision making (i.e., individuals make decisions individually instead). The comparison is *not* with an actual alternative governance arrangement for collective decision making like a dictator/manager who makes the collective decision on behalf of all in the society/organization). Our experiment addresses this gap and is the only one we are aware of whose specific focus is on how collective decisions from different governance arrangements affect organizational performance.

Section 2 reviews the related experimental literature. The experimental design and results follow in sections 3 and 4, respectively. Section 5 provides a discussion and conclusions. There is evidence of both the Galton wisdom of the crowd and Downs rational ignorance effects when all 'vote' on the collective decision as compared with when a 'dictator' makes it. There is virtually no evidence that 'voice' affects the quality of the collective decision via Milton's 'truth will out' or the specific additional incentive to process information. In contrast, the 'voice', and not the 'vote', aspect of governance matters for cooperativeness in the public goods decision. Indeed, somewhat surprisingly, 'voice' appears to have its most telling influence on motivation when combined with one person rather than all 'voting' on the collective decision.

## 2. Related experimental evidence

In relation to motivational change, several experiments compare the effect of punishment rules (and rewards) that are either given by nature or through a democratic voting procedure in what are public goods games (or real effort tasks) (e.g., see Feld and Tyran, 2002, Dal Bo et al, 2010, Sutter et al, 2010, and de Mellizo, Carpenter and Matthews, 2014). They find that

cooperation (or effort) rises when punishment rules are chosen democratically. Experiments that compare democratic collective decision making with an alternative of decentralization to individual decision making typically yield a similar conclusion (e.g. see Walker et al, 2000, and Hamman, Weber and Woon, 2011). However, there are experiments of this kind where there is no effect from voting by itself: e.g. Messer et al (2007) and Castillo et al (2018). Interestingly, Messer et al (2007) also considers the influence of 'voice' (i.e. chat), finding that it has a significant influence on contributions to a public good and that this is biggest when combined with voting.

There are some experiments that examine the influence of governance voting in other decision problems. For example, Sutter (2009) and Bornstein et al. (2004) respectively examine whether dictators or democracies are more inclined to take risks or swerve in chicken games. Brandts et al (2014) find that elected leaders are more effective than randomly selected ones in minimum effort games, but there are only two, to our knowledge, that are close to our concern with Mill's the motivational balance between individual and general interest that is revealed in public goods decisions. Frohlich et al. (1998) compare an employee owned and controlled (via democratic voting over remuneration scheme) firm with individual owned and managed firm using a real effort task. This combines differences in ownership with differences in the rules for collective decision making and so does not isolate the effect of the collective decision rule. It also does not focus on the possible shift from being motivated by individual interest to that of the general interest suggested by Mill. Nevertheless, their results are suggestive because productivity is higher with the employee owned and controlled firm and workers feel happier.

Ellman and Pezanis-Christou (2010) come closest to our paper because, unlike others, they focus on comparisons between rules for collective decision making. They compare the behavior of teams that are organized 'horizontally' and 'vertically' in making collective dictator game decisions. Their 'horizontal/vertical' distinction reflects whether all team

members or just one person make(s) the decision and so does capture the difference between democratic and hierarchical voting rules for decision making rules that interests us. They find that one type of horizontal team is more generous than others, including the 'vertical' organizations. However, they are interested in how the collective decisions from these different governance arrangements affect third parties and not, like us, in how they affect the performance of the organization.

Many experiments consider the effects of giving scope for 'voice'. Chat that precedes individual decision making in public goods, weakest link and gift exchange games typically improves outcomes (e.g. see Bochet et al 2006; Messer et al 2007, Sutter and Strassmair 2009; Cason et al. 2012 and Cooper and Lightle 2013); chat that precedes individual decision making improves the equitability of outcomes in a division game (Simon and Sulkin 2002), and may improve it in the context of an ultimatum game (Sulkin and Simon 2001). However, the comparison in these experiments is between chat and no chat without any background difference in the collective decision rules. Dickson et al (2007) do have chat that precedes voting, but there is no comparison with non-voting dictatorship or with the absence of chat. However, they do find evidence that subjects communicate not simply for strategic reasons but also for the free exchange of ideas.[4] Ellman and Pezanis-Christou (2010) do have this background and the effects of chat/'voice' can be disentangled. They find that communication is actually, and perhaps paradoxically, most powerful when combined with vertical (that is, hierarchical) collective decision making (albeit over what is a distribution type collective decision).

Several experiments examine wisdom of the crowd-like compositional effects of different governance rules. For example, Charness and Sutter (2012) contrast individual

[4] Dickson et al (2007) build on the theory model by Hafer and Landa (2003). Dickson et al. (2007), Simon and Sulkin (2002) and Sulkin and Simon (2001) belong to a tradition of empirical research in political science on the role of deliberation. Much of the other research in this tradition is based on polls and surveys rather than behavioral experiments; for a survey, see Karpowitz and Mendelberg (2011).

decisions with teams of individuals making the same decision collectively with and without chat. Teams are less prone to the errors in decision making that have been identified in many studies of individual decision making (e.g. with respect to Bayesian updating and the conjunction fallacy). Some evidence suggests that teams behave closer to standard game theory predictions because they develop a better understanding of the strategic aspects of an interaction (see Cooper and Kagel 2005). Again, though, this is not the comparison between democratic and hierarchical collective governance 'voting' rules that interests us. We are only aware of one paper that does this: Sutter (2009) contrasts individual risky investment decisions with those taken by a group under a democratic voting and under a circulating hierarchical rule. He finds no difference between the two collective decision rules.

A further compositional effect has been found in these experiments that may be relevant for the motivational hypothesis above. It concerns the pro-sociality of groups as compared with individuals. It seems that groups are less pro-social than individuals when making the same decision and this is possibly due to a 'responsibility-alleviation' effect whereby people feel less morally responsible for a joint decision than an individual one (see Charness and Sutter, 2012). While the contrast in these experiments is between groups and individuals and not between different ways of making a collective decision, the result seems to indicate that participation in decision making may reduce pro-sociality, contrary to Mill.

Lopez de Leon and Rizzi (2014) provide evidence of the Downs rational ignorance incentive effect in a natural experiment under democracy (and the fact that this is a natural experiment precludes the comparison with dictatorships). Against this, Hung and Plott (2001) provide evidence of an incentive effect in an information cascade experiment. They compare the outcomes when there is an individual reward system with one that rewards the correctness of the majority vote and find that performance is better under majority voting. What both miss is the contrast with hierarchical collective decision making. The experimental literature on the

decision to vote is also relevant for the Downs argument because it turns on the same logic concerning the chances of being the pivotal voter (e.g., see Levine and Palfrey 2007). The tests for an effect from this chance on voting are complicated by the possible expressive or partisan reasons for voting and by asymmetries in information (see Feddersen et al, 2009, and Battaglini et al, 2010). While often supportive of the importance of Downs's pivotal voter considerations, for obvious reasons, they do not involve a comparison with incentives faced by dictators.

In short, the question of whether dictators face any stronger incentives than electors to make an effort when making collective decisions, like that of whether there are compositional effects from this difference in how collective decisions are made, has not to our knowledge been addressed empirically. In the one case where the motivational effect has been considered across democracies and dictatorships, it has not focused on the cooperative decisions that interest us.

## 3. Experimental design

### *3.1 Overview*

The key trade-off in the design of the experiment is between simplicity and generality. We have favored simplicity: subjects make just two decisions. The first is a collective one. Individuals are organized into two groups that compete with each other over price. The price set by each group affects the revenue that comes to each group. We chose a price competition between different groups because competitions between organizations are common and the object of collective choice in this competition, a price, is a single, simple variable which has direct pay-off consequences for all. This creates a simple and clear incentive in the outcome of the collective decision. In addition, this collective decision is interestingly complicated in the sense that it is not a simple maximization problem as it involves judgment over the behavior of the other group. Further, and following from this, we have two simple measures of the

quality of the collective price decision: how far it deviates in any period from the best response to the price of the other group and by the total revenue secured over all time periods. The first captures how good any single group price decision or individual suggestion is. The virtue of the second is that, by looking at the effect over all pricing decisions, it allows for more complicated pricing strategies (e.g. where price in one period is set below the optimum for that period in order to encourage better prices in future periods) but it can only be applied to a group's prices decisions (and not individual price suggestions because there is no revenue variable associated with an individual suggestion).

The second is an individual decision and is independent of the collective price decision for the group: each individual makes a contribution to a public good that affects the group's costs of production. These individual decisions therefore, together with the collective price decision, determine individual pay-offs. Nevertheless, they are independent. This is important. Whatever happens as a result of the collective price decision, the individually self-interested public goods decision is to contribute 0. In so far as an individual contributes more than 0, then this can be used as an index of their pro-sociality towards their fellow group members (as is commonly the case in the interpretation of behavior in a public goods game).

The governance rules vary along two dimensions. Along the 'voting' dimension either every member of the group makes a suggestion for his/her group's price and the median suggestion is chosen for the group (=D, mimicking the median voter theorem) or one person from the group is selected and throughout the experiment this person decides on his/her group's price (=H). Although we used neutral terms in the experiment to describe these organizational differences, we call the H decision maker the dictator for convenience in what follows. We chose this random method for selecting the dictator because we wanted to focus on the pure effects of the difference in the mechanism for this collective decision. We did not wish this to be conflated with any possible selection effect whereby the dictator might either be resented or

accepted by others (or indeed was better or worse in the job on average than others) as a result of whatever merits, perceived or otherwise, they bought to the job in the first place.[5] Along the 'voice' dimension there is either scope for discussion within the group through an anonymous chat box (=Chat) or not.[6]

To test H1, given the above interpretation of behavior in the public goods game, we examine across the governance treatments whether the 'voting' and 'voice' rules affect contributions to the public good. We next test H2 by considering whether the typical individual suggestion for the price in a D group 'election' is worse than the price suggestion of the dictator/manager in the H group. Since the dictator/manager is selected at random, there should be no difference in his or her skills compared with the typical voter in a D group. Thus if the typical voter in D has a worse suggestion, then this points to a difference in information processing between a person in D and the dictator in H. Once we know the results of testing H2, we turn to H3. The quality of the D group's collective price decision can be compared with that of the H group can be determined in the same way using the deviation from best response prices for each group's governance arrangement. However, the interpretation of this evidence will depend on the outcome of H2 (and this is why we consider H2 first). This is because the Galton 'wisdom of the crowd' and Milton's 'truth will out' insights apply only when individuals in both D and H groups make the same efforts. Thus, for example, insofar as there is evidence to support H2, then a superior quality price in H than D need not count against H3 because the difference in quality could have arisen from the information processing differences. Likewise, if there is evidence for H2, but the quality of the prices decision is better

[5] For a similar reason we opted for equal individual shares in the revenue generated for the group by this price decision: we did not want to conflate the pure influence of the architecture of collective decision making with differences that might arise because either inequality varies or is perceived differently across these organizational forms. Again, this is not to suggest that different governance arrangements may not differ in this respect. They may well do.

[6] This set up closely follows van den Steen's (2009) model and seems to capture the essence of individual participation in many organizations/societies where free riding is a possibility over individual effort affecting the common good and where the organization/society must make collective decisions which affect its standing, and those who belong to it, when in competition with others.

than or the same in D as in H, then this would count in favor of H3 because despite the weaker information processing by people in D, the quality of the decision was not impaired compared with H.

*3.2 Experimental details*

At the beginning of each experimental session, 10 individuals are randomly allocated to two groups of 5. These groups are called 'organizations'. There are 20 periods in each session, where periods 1–10 and 11–20 are called stages 1 and 2 respectively and both decisions are made in each period.

The set-up for the price decision was presented as a symmetric revenue table in which there are two Pareto ranked pure Nash equilibria. The relationship between own organization's price, other organization's price and individual $i$'s revenue $r_i$ is given by Table 1. We employ a payoff table presentation to make the task as simple and transparent as possible for the subjects. [7] The efficient cooperative outcome is (80, 80) and the two pure Nash equilibria are (20, 20) and (30, 30). Consequently, the price decision is non-trivial in the sense that it does not simply turn on being able to read sensibly the revenue table associated with different price combinations. There are benefits from implicit cooperation between the two organizations if this can be achieved and, if this is not possible, it still pays to secure the Pareto superior Nash equilibrium (30, 30). We hope thereby to have built-in scope for judging the quality of collective decision making.

[7] We recognize that, by making the task easier to understand, a payoff table presentation might potentially increase collusion relative to a more opaque context, as found in Gürerk and Selten (2012). Future research could look at the impact of not having payoff table presentation, but, in a first experiment and given the overall complexity of the decision environment, we saw the facilitation of a clear understanding of the decision environment as the priority in terms of experimental inference. There is not in any case evidence of systematic collusion in our experiment (see the results section).

**Table 1 Usual Individual Revenue[8]**

| | | Price of other organization | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 |
| Price of own organization | 10 | 144 | 184 | 224 | 259 | 259 | 259 | 259 | 259 | 259 | 259 |
| | 20 | 188 | 268 | 348 | 428 | 482 | 482 | 482 | 482 | 482 | 482 |
| | 30 | 132 | 252 | 372 | 492 | 612 | 670 | 670 | 670 | 670 | 670 |
| | 40 | 0 | 136 | 296 | 456 | 616 | 776 | 821 | 821 | 821 | 821 |
| | 50 | 0 | 0 | 120 | 320 | 520 | 720 | 920 | 936 | 936 | 936 |
| | 60 | 0 | 0 | 0 | 84 | 324 | 564 | 804 | 1015 | 1015 | 1015 |
| | 70 | 0 | 0 | 0 | 0 | 28 | 308 | 588 | 868 | 1058 | 1058 |
| | 80 | 0 | 0 | 0 | 0 | 0 | 0 | 272 | 592 | 912 | 1066 |
| | 90 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 216 | 576 | 936 |
| | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 140 | 540 |

*Note*: table show the (equal) revenue for each of the participants in the own organization, depending on the price set by the own organization and the price set by that of the other, competing organization.

The set up for the second, individual, decision is a standard public goods game. Individuals in each organization make an individual decision ($x_i$) of contribution from an endowment of 100 points and this affects the individual costs of production ($c_i$) in their organization as in (1).

$$c_i = 160 + x_i - 0.4\sum_{j=1}^{5} x_j. \qquad (1)$$

It is 'as if' there is a certain amount of work that must be done and this can be undertaken more or less efficiently in terms of individual efforts depending on whether the individuals solve the free rider problem to capture the potential synergies that exist in production between these individual efforts. Individual earnings in the experiment are determined by the individual cost in (1) together with the equal share of the revenue that each individual receives as a result of the price decision.

Individual *i*'s pay-off in each period is thus

[8] This payoff table is applicable for all treatments except the first stage of *Shock* treatment.

$$\pi_i = r_i - c_i.$$

The prices of one's own and the other organization and the total contribution of one's own organization are given at the end of each period. Hence the individual does not know his/her revenue in the current period when he/she chooses his/her contribution to the public goods. But the outcomes in all previous periods are provided in the history table.

We have eight treatments. The first four (*Baseline, LimInfo, DDLimInfo* and *HHLimInfo)* are designed to check whether the nature of the information and the governance type of a group's competitor affects behavior.

***Baseline.*** The *Baseline* treatment has a D voting group over price competing with a an H group, where price is determined by a dictator, with the following information about the past history: they know their own and other's price and their own anonymous contributions to the public good, as well as the average payoff obtained in the past both by their own organization and the other organization.

***LimInfo, DDLimInfo* and *HHLimInfo***. These three treatments are control checks for whether any of the behavior observed in either type of group in the Baseline depends either on the fact that it is interacting with an organization of the other type or on whether each knows about how the other fares. Thus, the *LimInfo* treatment (for Limited Information) is the same as the Baseline except that each organization knows only their own average payoff and not that of the other organization (the other information is the same). Treatments *DDLimInfo* (for Democracy Democracy Limited Info) and *HHLimInfo* (for Hierarchical Hierarchical Limited Info) have the limited information as *LimInfo* but pitch, respectively, a D against another D and an H against an H (rather than a D versus a H).

The next four treatments expand inclusivity beyond the 'voting' dimension in the *Baseline* to include first the 'voice' dimension, then expand the domain of collective decision

making to an internal recommendation over the contribution to the public good and finally the introduction of 'voice'.

***ChatD.*** As in the *Baseline* treatment, there are a D voting and a H group competing with each other with knowledge of their own and other's price and of their own anonymous contributions to the public good, as well as of the average payoff obtained in the past both by their own organization and the other organization. The difference from the *Baseline* is that members of the D group *only* have now got a free form anonymous chat facility for the discussion of the price decision.[9]

***Rec.*** Treatment *Rec* (for Recommendation) is the same as the *Baseline* treatment except that the domain of discussion now explicitly expands to include a recommendation with respect to the individual contribution level for the public good. In D groups individuals vote over this recommendation and it is determined, as in the price decision, by the median individual vote. We also allow dictators in H groups to make a recommendation on how much to contribute to the public good because dictators or managers in hierarchic organizations can similarly communicate with their subjects or workers. Thus, this treatment preserves the voting distinction between D and H groups but allows internal recommendations to be made regarding the contribution to the public good using the same governance mechanism as the price (i.e. D or H respectively).

***ChatRec.*** Treatment *ChatRec* (for Chat and Recommendation) is the same as *Rec*, except that there is scope for discussion via an anonymous chat room facility in <u>both</u> D and H organization over both price and the public goods decisions.[10]

---

[9] Specifically, and as stated in the instructions, 'in alternate rounds, starting with the first round, each participant has the opportunity to discuss with others through written messages in an anonymous chat room what price should be set.' Subjects were not prevented from talking about contributions rather than just prices, but only occasionally did.

[10] Specifically, in addition to the text provided in ChatD (see previous footnote), in the section of the instructions dealing with contributions to the public good, subjects are told: 'In alternate rounds, starting with the first round, participants in both organisations have the opportunity to discuss with others in their organisation through written messages in an anonymous chat room what contribution should be made (at the same time as discussing what

**Table 2 Individual Revenue during the 1st Stage of the *Shock* Treatment**

| | | Price of other organization | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 |
| Price of own organization | 10 | 73 | 89 | 105 | 121 | 131 | 131 | 131 | 131 | 131 | 131 |
| | 20 | 106 | 138 | 170 | 202 | 234 | 248 | 248 | 248 | 248 | 248 |
| | 30 | 99 | 147 | 195 | 243 | 291 | 339 | 351 | 351 | 351 | 351 |
| | 40 | 52 | 116 | 210 | 288 | 368 | 448 | 518 | 518 | 518 | 518 |
| | 50 | 0 | 45 | 125 | 376 | 536 | 696 | 856 | 965 | 965 | 965 |
| | 60 | 0 | 0 | 30 | 264 | 504 | 744 | 984 | 1224 | 1339 | 1339 |
| | 70 | 0 | 0 | 0 | 0 | 272 | 592 | 912 | 1232 | 1552 | 1642 |
| | 80 | 0 | 0 | 0 | 0 | 0 | 240 | 640 | 1040 | 1440 | 1840 |
| | 90 | 0 | 0 | 0 | 0 | 0 | 0 | 168 | 648 | 1128 | 1608 |
| | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 56 | 616 | 1176 |

***Shock.*** This is identical to *ChatRec* except that the terms of the competitive interaction change in the first 10 periods. Specifically, in the first stage, the payoffs are such that the two Nash equilibria are now (50, 50) and (60, 60) and the cooperative outcome is (100, 100), as indicated by Table 2. The payoff table is the same as before in the second stage of 10 periods with (20, 20) and (30, 30) as Nash equilibria and (80, 80) as the cooperative outcome. The aim is to capture a positive demand shock to see if one organization is better able to adjust to a changing competitive environment.

The design examines changes across a variety of dimensions, and Table 3 provides a summary. It would have been impractical to explore all possible combinations of organizations in this multidimensional space. We have, however, ensured that there is always only a single change between pairs of treatments so that the comparison of behavior between treatments yields a clear interpretation. For example, by comparing behavior of D in *Baseline* and in *ChatD* we examine whether 'voice' makes a difference to a D organization. Likewise, by

price to set).' Unlike in ChatD, subjects regularly talked about both prices and contribution in ChatRec and Shock (described below).

comparing the behavior of D and H organizations in *Rec* and *ChatRec*, we test for the influence of 'voice' over both price and public goods decisions; and we test for the possible differential response to a shock between *ChatRec* and *Shock* because the only change is the shock to the pay-off matrix associated with the price decision. At the end of the experiment all subjects also performed a set of Holt and Laury (2002) type tasks.[11] Table 4 shows the sequence of tasks in each period.

**Table 3 Summary of Experimental Treatments**

| Treatment | *Baseline* | *LimInfo* | *DDLimInfo* | *HHLimInfo* |
|---|---|---|---|---|
| Organizations | D vs. H | D vs. H | D vs. D | H vs. H |
| The other organizations' mean payoff | Informed | | | |
| Chat on price decision | | | | |
| Chat on contribution decision | | | | |
| Recommendation over contribution | | | | |

| Treatment | *ChatD* | *Rec* | *ChatRec* | *Shock* |
|---|---|---|---|---|
| Organizations | D vs. H | D vs. H | D vs. H | D vs. H |
| The other organizations' mean payoff | Informed | Informed | Informed | Informed |
| Chat on price decision | D only | | D and H | D and H |
| Chat on contribution decision | | | D and H | D and H |
| Recommendation over contribution | | D and H | D and H | D and H |

**Table 4 The Sequence of Tasks in Each Period**

| Task | Treatment | Period |
|---|---|---|
| Chat on price decision | *ChatD* (D only), *ChatRec*, *Shock* | Every other period |
| Price decision | All treatments | Every period |
| Chat on contribution decision | *ChatRec*, *Shock* | Every other period |
| Recommend. over contribution | *Rec*, *ChatRec*, *Shock* | Every period |
| Contribution decision | All treatments | Every period |
| Feedback | All treatments | Every period |

[11] The tasks appeared on the screen in a randomized order. Each task consists of 10 questions to compare a pair of lotteries. Using these tasks we obtained measures of risk aversion, loss aversion and ambiguity aversion that we could use in the regression analysis. See the online appendix B for more details on the tasks.

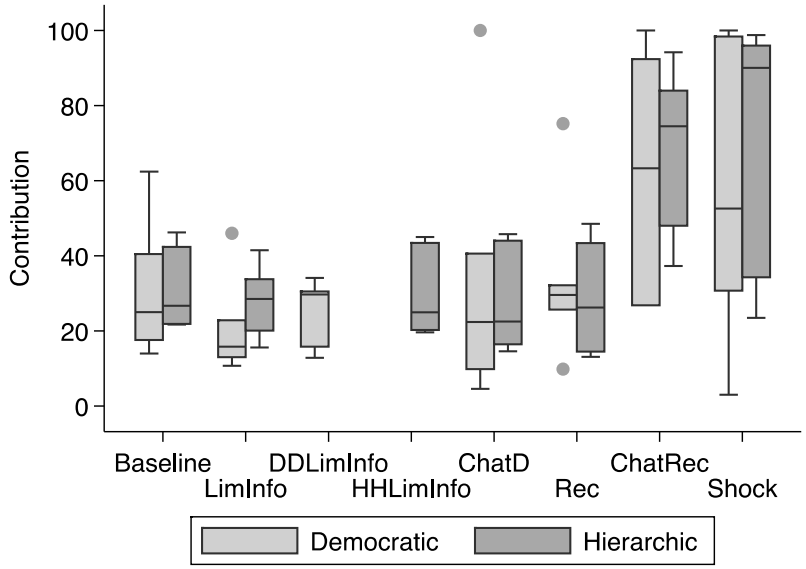
The experiment was conducted at the CBESS experimental laboratory at the University of East Anglia. Except for the instructions, the control questionnaire and the experiment were computerized with zTree (Fischbacher 2007). Each session consisted of 10 subjects, who were equally split into two groups. There were 7 sessions for each treatment, and so a total of 560 subjects were involved in 56 sessions. Almost all subjects were university students with various backgrounds. ORSEE (Greiner 2015) was used for recruitment. The experiment lasted between 90 minutes and two hours. The earnings were the sum of the profits in one randomly selected period from each stage, the total points in all four Holt and Laury type tasks and a show up fee of £2. Experimental points were translated into £s at a rate of £1 for 40 points. Subjects earned on average £11.25 in the main part and £3.00 in the end of experiment tasks, yielding a total average payment of £16.25.

## 4. Results

The results are summarized in the three series of box-plots in Figures 1-4 and in Table 5. Figure 1 gives the contribution to the public good in the second stage of each treatment by each type of organization, Figure 2 does the same for the price in the second stage; and Figures 3 and 4 give, respectively, revenue and payoffs from the second stage. In each session two groups were interacting and the average for each group is one data-point in these box-plots. We focus on stage 2 because it allows for any learning to have taken place but Table 5 provides the same information for periods 1-10 as well as the second stage of periods 11-20.[12]

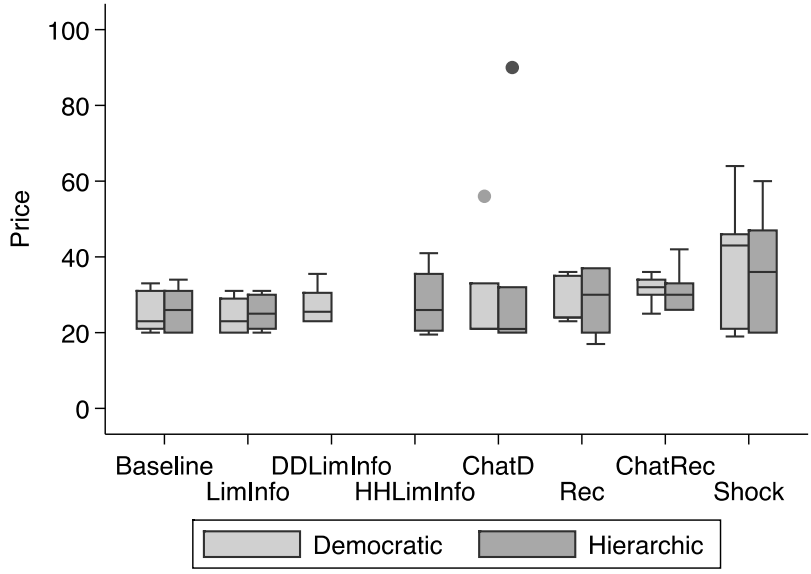
[12] See the online appendix C for period by period information on prices, contributions and payoffs. We find only weak end of game erosion in contributions. Prices are mainly around (20,20) and (30,30) and there is really only one session where both groups achieve the Pareto superior outcome of (80, 80) for half the periods. We could not in any obvious way connect the content of the chat in this session.

**Figure 1 Stage 2 Contribution to Public Good in D and H Organizations by Treatment**



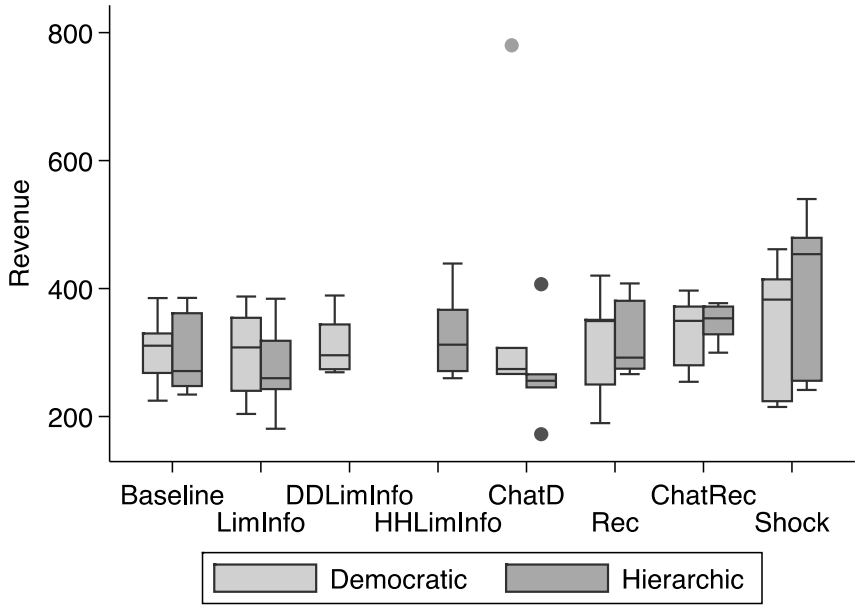

*Notes*: Each observation is the mean value at the session level. The middle bar refers to the median value; the edges of the box correspond to the 25th and 75th percentile; whiskers extend to 1.5 times the inter-quartile range; circles identify any other observation. Stage 2 corresponds to periods 11-20.

**Figure 2 Stage 2 Price Set in D and H Organizations by Treatment**



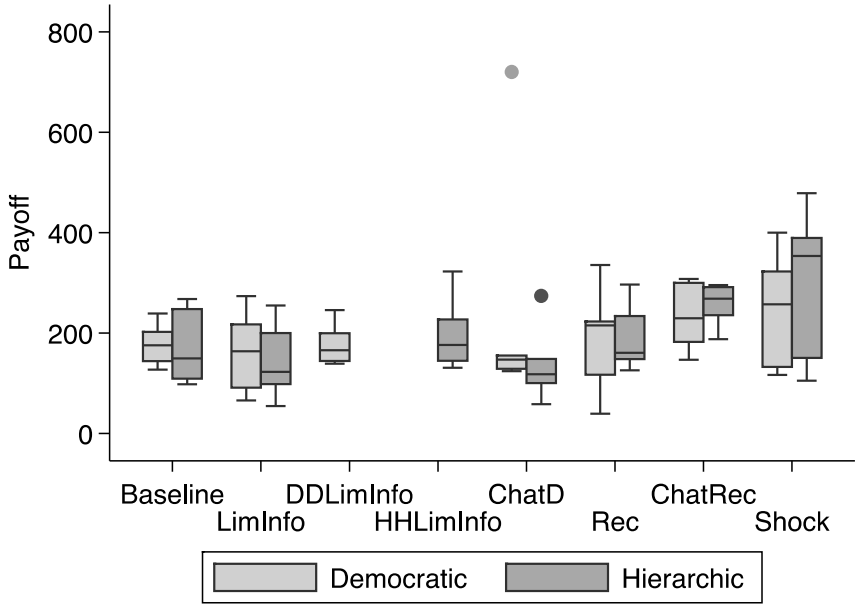

*Notes*: For the explanation, see the notes to Figure 1.

**Figure 3 Stage 2 Revenue in D and H Organizations by Treatment**




*Notes*: For the explanation, see the notes to Figure 1.

**Figure 4 Stage 2 Payoffs in D and H Organizations by Treatment**




*Notes*: For the explanation, see the notes to Figure 1.

**Table 5 Key Experimental Statistics**

| Treatment | Periods | Contribution | | | Price | | | Payoff | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | D | H | Total | D | H | Total | D | H | Total |
| *Baseline* | 1 – 10 | 34.3 | 31.7 | 33.0 | 34.2 | 29.3 | 33.4 | 210.8 | 171.8 | 191.3 |
| | 11 – 20 | 31.9 | 31.5 | 31.7 | 29.0 | 26.1 | 28.5 | 174.9 | 167.2 | 171.1 |
| | Overall | 33.1 | 31.6 | 32.4 | 31.6 | 27.7 | 30.9 | 192.8 | 169.5 | 181.2 |
| *LimInfo* | 1 – 10 | 25.6 | 32.4 | 29.0 | 32.4 | 31.6 | 32.3 | 199.7 | 189.0 | 194.4 |
| | 11 – 20 | 20.5 | 27.6 | 24.1 | 29.4 | 25.7 | 28.8 | 158.2 | 144.3 | 151.3 |
| | Overall | 23.0 | 30.0 | 26.5 | 30.9 | 28.6 | 30.5 | 179.0 | 166.7 | 172.8 |
| *DDLimInfo* | 1 – 10 | 30.7 | | 30.7 | 34.9 | | 34.9 | 225.2 | | 225.2 |
| | 11 – 20 | 24.2 | | 24.2 | 29.8 | | 29.8 | 174.3 | | 174.3 |
| | Overall | 27.4 | | 27.4 | 32.4 | | 32.4 | 199.8 | | 199.8 |
| *HHLimInfo* | 1 – 10 | | 33.3 | 33.3 | | 28.9 | 28.9 | | 188.6 | 188.6 |
| | 11 – 20 | | 30.2 | 30.2 | | 28.4 | 28.4 | | 197.9 | 197.9 |
| | Overall | | 31.7 | 31.7 | | 28.6 | 28.6 | | 193.2 | 193.2 |
| *ChatD* | 1 - 10 | 34.7 | 32.3 | 33.5 | 31.8 | 34.7 | 32.3 | 253.7 | 157.4 | 205.5 |
| | 11 - 20 | 33.7 | 27.5 | 30.6 | 30.0 | 32.4 | 30.4 | 224.6 | 132.8 | 178.7 |
| | Overall | 34.2 | 29.9 | 32.0 | 30.9 | 33.6 | 31.3 | 239.1 | 145.1 | 192.1 |
| *Rec* | 1 - 10 | 28.1 | 27.8 | 27.9 | 35.7 | 32.7 | 35.2 | 211.9 | 191.8 | 201.9 |
| | 11 - 20 | 32.9 | 27.7 | 30.3 | 31.9 | 27.9 | 31.2 | 183.4 | 191.5 | 187.4 |
| | Overall | 30.5 | 27.7 | 29.1 | 33.8 | 30.3 | 33.2 | 197.7 | 191.6 | 194.6 |
| *ChatRec* | 1 - 10 | 47.1 | 58.9 | 53.0 | 40.4 | 38.4 | 40.0 | 303.6 | 286.4 | 295.0 |
| | 11 - 20 | 61.6 | 69.8 | 65.7 | 33.3 | 31.0 | 32.9 | 232.7 | 259.2 | 245.9 |
| | Overall | 54.3 | 64.4 | 59.3 | 36.8 | 34.7 | 36.5 | 268.1 | 272.8 | 270.5 |
| *Shock* | 1 - 10 | 57.2 | 68.3 | 62.8 | 54.0 | 55.0 | 54.2 | 479.0 | 470.7 | 474.9 |
| | 11 - 20 | 56.5 | 70.0 | 63.2 | 39.8 | 37.0 | 39.3 | 235.9 | 304.9 | 270.4 |
| | Overall | 56.8 | 69.2 | 63.0 | 46.9 | 46.0 | 46.8 | 357.5 | 387.8 | 372.6 |

Three findings stand out:

1) Within each treatment, D organizations behave very similarly in all respects to H ones;

2) Between treatments, D organizations behave very similarly in all treatments except *ChatRec* and *Shock* (i.e. the two treatments where there is 'voice'); and likewise H organizations behave very similarly in all treatments except the same *ChatRec* and *Shock* where there is 'voice';

3) Both D and H organizations in *ChatRec* and *Shock* contribute more to the public good than in *Baseline*, *LimInfo*, *DDLimInfo*, *HHLimInfo*, *ChatD*, *Rec* and this feeds through to higher payoffs in these treatments.

**Table 6 Wilcoxon Signed-Rank Tests for Difference between Democratic (D) and Hierarchical (H) Organizations in Stage 2 by Treatment**

| Treatment | Contribution | Price Level | Best Price Response | Revenue | Payoff |
|---|---|---|---|---|---|
| *Baseline* | 0.735 | 0.497 | 0.343 | 0.799 | 0.866 |
| *LimInfo* | 0.028[a] | 0.734 | 0.931 | 0.310 | 0.398 |
| *ChatD* | 0.499 | 0.660 | 0.498 | 0.176 | 0.310 |
| *Rec* | 1.000 | 0.735 | 0.665 | 0.612 | 0.735 |
| *ChatRec* | 0.499 | 0.796 | 0.476 | 0.553 | 0.091[a] |
| *Shock* | 0.612 | 0.496 | 0.372 | 0.176 | 0.091[a] |

*Notes*: Each observation is the mean value at the session level. [a] H value > D value.

**Table 7 Wilcoxon Rank-Sum Test for Treatment Differences in Relation to Democratic and Hierarchical Organizations in Stage 2**

| | Hierarchical Organizations | | | | | |
|---|---|---|---|---|---|---|
| | *Baseline* vs. *LimInfo* | *LimInfo* vs. *HHLimInfo* | *Baseline* vs. *ChatD* | *Baseline* vs. *Rec* | *Rec* vs. *ChatRec* | *ChatRec* vs. *Shock* |
| Contribution | 0.482 | 0.565 | 0.565 | 0.565 | 0.006[b] | 0.655 |
| Price Level | 0.898 | 0.654 | 0.647 | 0.653 | 0.479 | 0.305 |
| Best Price Response | 0.743 | 0.336 | 0.475 | 0.555 | 0.394 | 0.647 |
| Revenue | 0.565 | 0.142 | 0.225 | 0.277 | 0.406 | 0.482 |
| Payoff | 0.565 | 0.110 | 0.277 | 0.482 | 0.048[b] | 0.338 |
| | Democratic Organizations | | | | | |
| | *Baseline* vs. *Liminfo* | *LimInfo* vs. *DDLimInfo* | *Baseline* vs. *ChatD* | *Baseline* vs. *Rec* | *Rec* vs. *ChatRec* | *ChatRec* vs. *Shock* |
| Contribution | 0.085 [c] | 0.338 | 0.655 | 0.749 | 0.110 | 0.898 |
| Price Level | 0.797 | 0.335 | 0.844 | 0.139 | 0.334 | 0.482 |
| Best Price Response | 0.745 | 0.197 | 0.699 | 0.896 | 0.562 | 0.364 |
| Revenue | 0.848 | 0.655 | 0.404 | 0.749 | 0.443 | 0.655 |
| Payoff | 0.565 | 0.655 | 0.406 | 0.848 | 0.225 | 0.949 |

*Notes*: Each observation is the mean value at the session level. [b] *ChatRec* value > *Rec* value. [c] *Baseline* value > *LimInfo* value.

Tables 6 and 7 largely support these findings. On 1), Table 6 gives the *p* values for the Wilcoxon signed-rank test for differences between D and H in each treatment (i.e. finding 1).[13] The only statistically significant difference is with respect to contributions in *Liminfo,* but, as we note below, this difference is not found in the individual regressions.[14] On 2) and 3),

Table 7 gives the *p* values for the Wilcoxon rank-sum test for differences in D behavior and differences in H behavior between comparable treatments: no *p* value is significant at the 5% level except between *Rec* and *ChatRec* for contributions ($p < 0.01$) and correspondingly for payoffs ($p = 0.05$).

We now turn to individual behavior. Table 8 reports the regressions on individual contributions to the public good using multilevel mixed-effects estimation technique (GLS random-effect linear regressions produce similar results and are reported in the electronic Appendix). We control for possible differences in competitor type and information, whether there is chat over price and chat over price and contribution, if the person is the dictator, the voting rule for joint decisions, the actual contribution recommendation (when applicable), time trend and some demographic variables. Table 9 does the same for the quality of the price decision in H and price suggestions in D, except Regressions 19 and 26, where the resulting price quality of the organization is regressed. Quality here is measured by the absolute difference of the individual decision/suggestion from the best response price; a lower value corresponds to a better price quality. There is evidence of a standard progressive decline in

[13] Tables 5 and 6 report the Stage 2 results. See the online appendix D for the Stage 1 results.

[14] All reported *p* values in this paper are two tailed. One tailed tests are available in the online Appendix. We prefer two tailed tests here because they permit a comparison for differences (e.g. over price) where our hypotheses have no clear implication. In addition, although our hypotheses are framed in terms of, for example, inclusiveness encouarging cooperation, we are also interested in the possibility that it might discourage cooperation (and not just whether it encourages cooperation or fails to). For this reason, the difference from zero is more interesting than the difference being greater than zero.

contributions in Table 8, and of an equally plausible progressive improvement in the quality of price decisions in Table 9. We now move to our key results.

**Table 8 Regressions on Stage 2 Individual Contributions**

| Regression | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Organization | H | D | All | H | D | All |
| Not Mixed | 1.051 | 3.643 | 2.548 | 2.620 | 3.671 | 3.146 |
| | (8.163) | (12.48) | (6.758) | (8.264) | (12.13) | (6.705) |
| No Payoff Info | -1.103 | -12.00 | -4.452 | -1.951 | -11.35 | -4.761 |
| | (7.327) | (12.66) | (6.885) | (7.410) | (12.31) | (6.836) |
| Chat Price | | 0.924 | 10.04* | | 1.803 | 9.358* |
| | | (12.68) | (5.288) | | (12.31) | (5.325) |
| Chat Contribution | | 29.08 | 26.09*** | | 26.82 | 25.99*** |
| | | (17.94) | (9.442) | | (17.41) | (9.396) |
| Chat Price and Contribution | 41.39*** | | | 42.09*** | | |
| | (8.540) | | | (8.556) | | |
| Recommend. Of Contribution | -1.717 | 1.198 | 2.331 | -1.820 | 1.071 | 1.515 |
| | (7.333) | (12.67) | (6.895) | (7.410) | (12.31) | (6.836) |
| Shock | 1.472 | -6.213 | -2.774 | 0.191 | -5.089 | -2.449 |
| | (8.562) | (12.66) | (7.848) | (8.556) | (12.31) | (7.742) |
| Dictator | -7.193*** | | -6.828** | -6.701** | | -6.701** |
| | (2.781) | | (3.284) | (2.772) | | (3.258) |
| Democracy | | | -7.009*** | | | -5.869** |
| | | | (2.300) | | | (2.306) |
| Round | -0.701*** | -0.389*** | -0.543*** | -0.698*** | -0.346*** | -0.522*** |
| | (0.130) | (0.113) | (0.0860) | (0.128) | (0.115) | (0.0858) |
| Age | -0.191 | 0.236 | -0.0163 | | | |
| | (0.284) | (0.274) | (0.231) | | | |
| Gender | 3.405 | 1.675 | 3.339* | | | |
| | (2.434) | (2.315) | (1.954) | | | |
| UK | -6.246** | 0.930 | -3.706 | | | |
| | (2.899) | (2.753) | (2.321) | | | |
| China | -7.264 | 3.753 | -6.453 | | | |
| | (5.096) | (4.735) | (4.021) | | | |
| Economics | -7.696** | -5.723 | -6.770** | | | |
| | (3.578) | (3.585) | (2.940) | | | |
| Business | -3.283 | -2.158 | 0.116 | | | |
| | (3.719) | (3.320) | (2.865) | | | |
| Loss Aversion | 0.860 | 0.709 | 0.999** | | | |
| | (0.576) | (0.560) | (0.457) | | | |
| Ambiguity Aversion | -0.392 | 0.339 | -0.314 | | | |
| | (0.362) | (0.352) | (0.294) | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Risk Aversion | -2.034** | -1.390** | -1.899*** | | | |
| | (0.833) | (0.674) | (0.601) | | | |
| Constant | 47.63*** | 27.46** | 39.74*** | 34.72*** | 33.77*** | 35.29*** |
| | (10.05) | (12.29) | (8.424) | (4.371) | (8.730) | (4.187) |
| Observations | 2,730 | 2,790 | 5,520 | 2,800 | 2,800 | 5,600 |

*Notes*: The regressions are three-level mixed models with random intercepts at both the session and the subject-within-session levels; Regressions 1and 4 include data only from Hierarchy organizations, 2 and 5 only from Democratic organizations, and 3 and 6 from both, respectively. Variables: Not Mixed: only organizations with the same structure in the treatment (*DDLimInfo* and *HHLimInfo* treatments); No Payoff Info: organizations get no payoff info of the other organization (*DDLimInfo*, *HHLimInfo*, *LimInfo*); Chat Price: organization can chat about price (treatment *ChatD* for D, *ChatRec* and *Shock* for both D and H); Chat Contribution: organization can chat about contribution (*ChatRec*, *Shock* for both D and H); Chat Price and Contribution: H organizations have either no chat or chat about both price and contribution; Recommendation of Contribution: D organization or CEO/dictator recommends a contribution (*Rec*, *ChatRec*, *Shock*); Shock: the *Shock* treatment in which there was a doubling of the revenue in the first 10 periods; Round: round number; Dictator = 1 when participant was a CEO (i.e. set price unilaterally); Democracy = 1 for democratic organizations; Gender = 1 for female; UK 1= from the UK; China 1 = from China; Economics 1 = Economics; Business 1 = Business studies; Loss, Ambiguity and Risk Aversion (from end of experiment tasks).

**Table 9 Regressions on Stage 2 Price Decision/Suggestion Quality**

| Regression | (7) | (8) | (9) | (10) | (11) | (12) | (13) |
|---|---|---|---|---|---|---|---|
| Organization | H | D | All | H | D | All | All (Org.) |
| Not Mixed | -4.925 | -0.0165 | -1.205 | -2.286 | 0.200 | -0.747 | -0.857 |
| | (3.519) | (3.079) | (2.832) | (3.268) | (3.022) | (2.815) | (2.763) |
| No Payoff Info | -0.284 | 2.281 | 0.802 | -1.214 | 1.457 | 0.287 | -1.473 |
| | (3.292) | (3.298) | (2.993) | (3.268) | (3.240) | (2.973) | (2.793) |
| Chat Price | | -1.230 | -3.573 | | -2 | -3.739 | -3.608** |
| | | (3.320) | (2.791) | | (3.240) | (2.767) | (1.642) |
| Chat | | -2.243 | 0.816 | | -1.257 | 1.192 | 4.751 |
| | | (4.700) | (4.254) | (4.582) | (4.204) | (3.588) | (3.604) |
| Chat Price and | 2.198 | | | 1.000 | | | |
| | (4.244) | | | (3.774) | | | |
| Recommend. of | -1.873 | 2.749 | 0.833 | -1.929 | 1.229 | -0.0223 | -1.831 |
| Contribution | (3.565) | (3.306) | (2.996) | (3.268) | (3.240) | (2.973) | (2.793) |
| Shock | -0.844 | 3.126 | 2.849 | 1.714 | 3.057 | 2.833 | 2.500 |
| | (4.322) | (3.299) | (3.208) | (3.774) | (3.240) | (3.164) | (3.190) |
| Democracy | | | 6.082*** | | | 6.143*** | 0.248 |
| | | | (1.497) | | | (1.462) | (0.677) |
| Age | -0.743*** | -0.467*** | -0.511*** | -0.784*** | -0.467*** | -0.520*** | -0.687*** |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | (0.150) | (0.0779) | (0.0697) | (0.145) | (0.0777) | (0.0691) | (0.0958) |
| Gender | -0.441 | 0.205 | 0.173 | | | | |
| | (0.272) | (0.144) | (0.126) | | | | |
| UK | -1.766 | -1.268 | -1.010 | | | | |
| | (2.090) | (1.237) | (1.056) | | | | |
| China | -4.131* | -0.290 | -0.446 | | | | |
| | (2.400) | (1.443) | (1.239) | | | | |
| Economics | -7.151 | -1.402 | -1.180 | | | | |
| | (4.902) | (2.508) | (2.175) | | | | |
| Business | -0.395 | -0.346 | -0.594 | | | | |
| | (3.007) | (1.911) | (1.619) | | | | |
| Loss Aversion | -4.491 | -0.646 | -0.912 | | | | |
| | (3.222) | (1.778) | (1.521) | | | | |
| Ambiguity | -0.957 | 0.343 | 0.267 | | | | |
| | (0.707) | (0.298) | (0.262) | | | | |
| Risk Aversion | -0.409 | -0.340* | -0.318** | | | | |
| | (0.290) | (0.187) | (0.158) | | | | |
| Constant | 0.397 | -0.623* | -0.500* | | | | |
| | (0.630) | (0.355) | (0.302) | | | | |
| Observations | 530 | 2,790 | 3,320 | 560 | 2,800 | 3,360 | 1,120 |

*Notes*: The dependent variable is the quality of the price vote suggestions, except for Regression 13, where it is the resulting price quality of the organization. Regressions 7 and 10 are two-level mixed models with random intercepts at the subject level only since there is only one dictator in each session. Regressions 8, 9, 11, 12 and 13 are instead three-level mixed models with random intercepts at both the session and the subject-within-session levels, except Regression 13 that has the organization-within-session at the nested-level error. For the description of the explanatory variables, see Table 8.

**Result 1**: There is no evidence that the behavior of D or H organizations depends on the organizational type of its competitor or information regarding the competitor's payoffs.

**Support**. In Table 8, the dummy variables for organizational type of the competitor (*Not Mixed*), for information on the competitor's payoffs (*No Payoff Info*) and for the interaction between D and No Payoff Info are not statistically significant in any of the regressions. The Wilcoxon tests of Table 6 also largely support this.

**Result 2**: Against H1a, there is no evidence that inclusiveness via 'voting' leads to greater cooperation in public goods interactions in D relative to H organizations. If anything, individuals making the price decisions in both D (i.e. everyone) and H contribute less to the public good than the non-decision makers in H. D and H organizations also earn similar payoffs.

**Support**. There is no evidence of differences in aggregate public good contribution between D and H organizations in Table 6. Tables 8 on individual contributions shows that the coefficients on *Democracy* are never significantly positive: the coefficient on Democracy is significant in regression 3 ($p < 0.05$) and in regression 6 ($p<0.10$) and negative suggesting that, if anything, contributions are worse under Democracy. The coefficient on the individual Dictator is also negative ($p < 0.05$ in regressions 3, 4, 6 and $p < 0.01$ in regression 1). The Democracy and Dictator coefficients are not significantly different from each other (the $p$ values are 0.958 and 0.808 for regressions 3 and 6 respectively). Thus the Dictator and the democratic decision makers appear to contribute less than the non-decision makers in H.

**Result 3**: In support of H1b, there is evidence for a positive effect of inclusiveness via 'voice' on public good contribution, particularly in H organizations.

**Support**. As noted earlier, Figure 1 is suggestive of this result. A nonparametric test in Table 7 finds the difference between *Rec* and *ChatRec* significant for H organizations ($p = 0.006$). Table 8 also finds that chat over price and contributions have a significant effect for H organizations (regressions 1 and 4) and chat over contribution is significant for the overall sample (regressions 3 and 6).

**Result 4a:** In support of H2a, inclusiveness via 'voting' discourages effort in the processing of information in D organizations.

**Result 4b:** Against H2b, there is no evidence that 'voice' affects the quality of individual price suggestions.

**Support.** In regressions 9 and 12 on the quality of individual price suggestions,[15] the Democracy dummy captures the price quality of the suggestions by individual decision makers in D relative to Dictators in H. It is positive and statistically significant ($p < 0.001$), indicating D proposers make worse suggestions than the Dictators. Since the dictator is chosen at random, there should be no difference in skill between the typical dictator and typical members of a D organization, and so it would seem that dictators must apply themselves more diligently to the problem of price setting than the typical member of a D organization. This is evidence of incentive effects under H2a. Against H2b, none of the chat coefficients are significant for the individual regressions Table 9 (regressions 7-12).

**Result 5a:** D and H organizations do not differ in terms of the quality of price decisions. Given that there is support for better quality of price suggestions in H organizations (Result 4a), the absence of a difference in the quality of the price decision between D and H must be due to a compositional effect, implying support for H3a.

**Result 5b:** Against H3b, there is no evidence that 'voice' affects the quality of price decisions. This is because there is no difference in the quality of price suggestions (Result 4b) and, in the light of this, the lack of a difference between D and H organizations for the quality of price decisions implies the lack of a composition effect.

[15] This means that non dictators are excluded since they did not make any price decision.

**Support**. Table 6 shows no difference between D and H in overall quality of price decisions in both the absolute difference from best response measure of quality and the average revenue one. Regression 13 in Table 9 confirms this. These regressions use the decision of each organization – whether achieved democratically or hierarchically – as an observation. In relation to this regression, the Democracy dummy captures whether D organizations made better or worse price decisions than H organizations: the coefficient is not statistically significant. Given Result 4a, this supports H3a.

Against H3b and given Result 4b, there are no differences in Table 7 for 'best response price' and 'revenue' between *Baseline* versus *Chat* for D and between *Rec* versus *RecChat* for D and H; and the only Chat coefficient that is significant is in regression 13, Chat price.

## 5. Discussion and Conclusions

Result 2 is important. It counts against the Mill/De Tocqueville motivational conjecture that the involvement through 'voting' on collective decisions under representative government encourages the application of principles that promote the general good. Indeed, the evidence that the decision makers in both D and H organizations contribute less to the public good suggests the reverse. It is 'as if' individuals had a fixed amount of public spiritedness and some gets used up in contributing to collective price decision with the result that less is available for the public good. While this is only a conjecture, it is consistent with moral licensing research where past moral behavior helps remove the concern of appearing uncaring in subsequent moral tasks (e.g., Mazar and Zhong, 2010; Mullet and Monin, 2016).

Result 2 is particularly important in the context of the experimental literature that contrasts the determination of collective rules by democratic voting and their determination by 'nature' because this typically suggests that there are public spirit benefits from democratic

voting. Result 2 cautions against generalizing from this literature to what is different choice between democratic and hierarchical voting rules for making collective decisions. In this respect, it is closer in spirit to the finding in that literature of Messer et al (2007).

Result 3, however, suggests that the 'voice' aspect of inclusiveness does have the Hirschman conjectured positive motivational effect on pro-sociality, something which in itself is not surprising given the evidence on chat from other experimental evidence. What is surprising is that it is strongest when combined with a hierarchical determination of collective decisions. This result is consistent with the experimental results on the effect of chat in public goods games when there is no institutional backdrop of either voting or dictatorship for collective decision making. The specific strength on motivation under the hierarchical determination of collective decisions is perhaps not so surprising given some models of deliberation (e.g. Hafer and Landa 2003) and it is also consistent with the one experiment where there is a version of this institutional backdrop: Ellman and Pezanis-Christou (2010) report a similar pro-social effect for contributions in dictator games.

The specific strength of 'voice' on motivation under hierarchical price setting decision making is at first sight puzzling, but our experiment does suggest one possible reason. The difference with 'voting' over collective decisions emerges in our experiment when there are recommendations over contributions to the public good that either are or are not accompanied by chat. It is possible that the same sense of a constant amount of public spiritedness, alluded to above, applies to these recommendations. As a result, with more individuals involved in the recommendation in a D than H organization, the motivational force of 'voice' appears to be correspondingly weaker in D than H. But why should the act of recommending or voting over recommendations weaken public spiritedness in our experiment? One explanation turns on the way that the small likelihood of being the pivotal voter in an election diminishes the incentive to vote (as well acquire and process information). Thus, although our subjects belong to small

electorates, it is possible that they may not be inclined to vote. That they are required to vote in the experiment could, therefore, be seen as a chore and this may explain why the act of voting seems to use up public spiritedness in our experiment. The same logic, though, does not seem to apply in our experiment to the 'voice' aspect of inclusivity because it promotes cooperation (Result 3). Again incentives may explain this difference. While there may be weak incentives to vote because of the small chance of being the pivotal voter, voice need not suffer from this small marginal effect because it is possible through persuasion that a single individual can have a large effect. As a result, people may feel very differently about participating in discussion as compared with participation through voting. In short, 'voting' may be a chore, while having scope for 'voice' may be perceived as something worthwhile.[16]

Taken together, Results 2 and 3 suggest that it is important to distinguish between the two dimensions of inclusivity: the 'voice' aspect of inclusiveness matters for motivation, but the voting aspect does not. Results 4 and 5, reinforce this point, but for the opposite reason. They suggest in contrast that the 'voting' and not 'voice' matters for both the wisdom of the crowd-like and the rational ignorance effects that follow from 'voting'. That the two effects seem to offset each other in terms of the resultant quality of collective decisions may be an artifact of our experiment. The fact that they work in opposite directions is what is important. Together with the evidence that 'voting' does not appear to matter for motivation, this may explain why, in an evolutionary context, we find examples of both voting and hierarchical types of organization. The performance of D and H organizations is largely the same in our experiment. The decisive difference for performance is the presence of 'voice' and this attribute of inclusivity can be found in D and H organizations (of course, this is consistent with the argument in Mukand and Rodrik (2015).

[16] This difference may explain why participation in elections has been falling and why it might be a mistake to infer from this that the commitment to inclusiveness *in the round* is waning.

Our results, particularly Result 2, could owe something to the economic frame of the two decision problems. Although the groups are referred to as organizations, they compete over price and this is likely to suggest a business organization where hierarchical forms of decision making are commonly encountered. As a result, hierarchical forms of decision making may be regarded as natural in this setting. This was a design decision. There are major benefits, particularly in relation to internal validity, from simplicity in experiments. We wanted the price suggestions decisions to have clear consequences for each individual and we wanted another decision to be related to but independent of the first decision where motivational changes could thereby be tracked independently of the quality of the collective decision. The price and cost setting does this. Having such a contextual frame is natural in terms of our incentive structure and, as in other settings (Alekseev et al 2017), it is likely to facilitate understanding of the experimental instructions. Nevertheless, even if we were consider the conclusion as being restricted to these economic domains, it is important.[17]

We recognize that the evidence from LimInf is only suggestive that the identity of the competitor does not matter in other information settings, and it is left to future research to investigate this further.[18] Another possible interesting avenue for future research would be to see whether one-sided messaging would be sufficient to operate as a mechanism 'voice', as opposed to the more natural chat setting that we employ.

Overall, we find that inclusivity in decision making matters for performance but not always in the ways that one might expect from most of the experimental literature that compares democratic collective decisions with either their determination instead by nature or

[17] This is to take a literal view of external validity in the sense that the results apply only to situations in so far as they correspond directly to those in the experiment. An alternative view holds that external validity comes because experiments are well suited to tracking the habits of decision making that subjects use in the world and bring to the lab.

[18] We would have thought to check for competitor identity effects in the other treatments had there been any reason to suppose the difference between LimInf and the Baseline (i.e. knowing the average pay-off of the competitor) was likely to interact with the type of competitor to make a difference.

through devolution to individual decision making. When democratic voting is contrasted with dictatorship for collective decision making, if anything, voting has a negative motivational effect on the willingness to cooperate. The key dimension of inclusivity that boosts cooperation is 'voice'. This is so in our experiment for organizations that make collective decisions through 'voting' and for those that do so hierarchically. This does not mean that 'voting' aspect of inclusivity does not matter. It does. 'Voting', indeed more than 'voice', is the key dimension affecting the processing and aggregating information in the organizations in our experiment. Nevertheless, the decisive source of difference in performance in our experiment comes from whether or not there is 'voice'.

**References**

Acemoglu, Daren, Simon Johnson, James A. Robinson, and Pierre Yared. 2008. "Income and Democracy." *American Economic Review* 98 (3): 808-42.

Alekseev, Aleksandr, Gary Charness and Uri Gneezy. 2017. "Experimental methods: When and why contextual instructions are important." *Journal of Economic Behavior and Organization* 134: 48-59.

Battaglini, Marco, Rebecca B. Morton and Thomas R. Palfrey. 2010. "The Swing Voter's Curse in the Laboratory." *Review of Economic Studies* 77 (1): 61-89

Black, Sandra E., and Lisa M. Lynch. 2004. "What's Driving the New Economy: The Benefits of Workplace Innovation." *Economic Journal* 114 (93): 97-116.

Bochet, Olivier, Talbot Page, and Louis Putterman. 2006. "Communication and punishment in voluntary contribution experiments." *Journal of Economic Behavior and Organization* 60 (1): 11-26.

Bornstein, Gary, Arthur Schram and Joep Sonnemans. 2004. "Do democracies breed chickens." In *Contemporary Psychological Research on Social Dilemmas*, R. Suleiman, D. Budescu, I. Fischer and D. Messick (eds.), Cambridge University Press, pp. 248-268.

Bowles, Samuel, and Herbert Gintis. 1987. *Democracy and Capitalism*. New York: Basic Books.

Bowles, Samuel, and Herbert Gintis. 1993. "Agency Theory and the Democratic Firm." In S. Bowles, H. Gintis and B. Gustaffson (eds), *Markets and Democracy.* Cambridge, Cambridge University Press.

Brandts, Jordi, David Cooper and Roberto Weber. 2014. "Legitimacy, Communication and Leadership in the Turnaround Game." Management Science 61(11): 2627-2645.

Cason, Timothy N., Roman M. Sheremeta, and Jingjing Zhang. 2012. "Communication and efficiency in competitive coordination games." *Games and Economic Behavior* 76 (1): 26-43

Castillo, Jose, Zhicheng Xu, Ping Zhang and Xianchen Zhu. 2018. "The effects of centralized power and institutional legitimacy on collective action." Working Paper, SSRN (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3127775).

Charness, Gary. 2000. "Responsibility and effort in an experimental labor market." *Journal of Economic Behavior and Organization* 42 (3): 375-384.

Charness, Gary, and Matthias Sutter, M. 2012. "Groups make better self-interested decisions." *Journal of Economic Perspectives* 26 (3): 157-176.

Cooper, David J., and John H. Kagel. 2005. "Are Two Heads Better Than One? Team Versus Individual Play in Signaling Games." *American Economic Review* 95 (3): 477–509.

Cooper, David J., and John P. Lightle. 2013. "The gift of advice: communication in a bilateral gift exchange game." *Experimental Economics* 16 (4): 443-477.

Dal Bó, Pedro, Andrew Foster, and Louis Putterman. 2010. "Institutions and Behavior: Experimental Evidence on the Effects of Democracy." *American Economic Review* 100 (5): 2205–2229.

Degan, Arianna. 2006. "Policy Positions, Information Acquisition and Turnout." *Scandinavian Journal of Economics* 108 (4): 669–82.

Dickson, Eric, Catherine Hafer and Dimitri Landa. 2008. "Cognition and strategy: a deliberation experiment." *Journal of Politics* 70(4): 974-989.

De Tocqueville, Alexis. 1840. *Democracy in America, Volume 2.* iBook.

Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper and Row.

Dryzek, John and Christian List. 2003. "Social choice and deliberative democracy: a reconciliation." *British Journal of Political Science* 33(1): 1-28.

Ellman, Matthew, and Paul Pezanis-Christou. 2010. "Organizational structure, communication and group ethics." *American Economic Review* 100 (5): 2478-2491.

Falk, Armin, and Nora Szech. 2013. "Morals and the market." *Science* 340 (6133): 707-711.

Feddersen, Timothy, and Alvaro Sandroni. 2006. "Ethical Voters and Costly Information Acquisition." *Quarterly Journal of Political Science* 1 (3): 287–311.

Feddersen, Timothy, Sean Gailmard, and Alvaro Sandroni. 2009. “Moral Bias in Large Elections: Theory and Experimental Evidence.” *American Political Science Review 103*(2): 175-192.

Feld, Lars and Jean-Robert Tyran. 2002. “Tax evasion and voting: an experimental analysis”, *Kyklos*, 55, 197-221.

Fischbacher, Urs. 2007. “z-Tree: Zurich Toolbox for Ready-made Economic Experiments.” *Experimental Economics* 10 (2): 171-178.

Frey, Bruno S. 1998. “Institutions and Morale: The Crowding-out Effect.” In *Economics, Values, and Organization*, ed. A. Ben-Ner and L. Putterman, 437–60. New York: Cambridge University Press.

Frohlich, Norman, John Godard, Joe A. Oppenheimer, and Frederick A. Starke. 1998. “Employee Versus Conventionally-Owned and Controlled Firms: An Experimental Analysis.” *Managerial and Decision Economics* 19 (4-5): 311–326.

Galton, Francis. “Vox populi (the wisdom of crowds)." *Nature* 75 (1907): 450-451.

Glaeser, Edward, La Porta, Rafael, Lopez de Silanes, Florencio and Shleifer, Andrei. 2004. “Do Institutions Cause Growth?” *Journal of Economic Growth* 9(3): 271–303.

Greiner, Ben. 2015. “Subject pool recruitment procedures: Organizing experiments with ORSEE.” *Journal of the Economic Science Association* 1 (1): 114-125.

Gürerk, Özgür, and Selten, Reinhard. 2012. “The effect of payoff tables on experimental oligopoly behavior.” *Experimental Economics* 15(3), 499-509.

Hafer, Catherine and Dimitri Landa. 2003. “Deliberation as self discovery and institutions for political speech” *Journal of Theoretical Politics* 19(3): 329-360.

Hamman, John, Roberto Weber, and Jonathan Woon. 2011. “An experimental investigation of electoral delegation and he the provision of public goods.” *American Journal of Political Science* 55(4): 738-752.

Helliwell, John. 1994. “Empirical Linkages Between Democracy and Economic Growth.” *British Journal of Political Science* 24 (02): 225-248.

Hirschman, Albert. 1970. *Exit, Voice and Loyalty*. Cambridge: Harvard University Press.

Holt, Charles A., and Susan K. Laury. 2002. “Risk Aversion and Incentive Effects.” *American Economic Review* 92 (5): 1644–1655.

Hung, Angela, and Charles Plott. 2001. “Information Cascades: Replication and an Extension to Majority Rule and Conformity-Rewarding Institutions.” *American Economic Review* 91(5): 1508-1520.

Ichniowski, Casey, and Kathryn Shaw. 2003. “Beyond Incentive Pay: Insiders' Estimates of the Value of Complementary Human Resource Management Practices.” *Journal of Economic Perspectives* 17 (1): 155–180.

Karpowitz, Christopher F., and Tali Mendelburg. 2011. “An Experimental Approach to Citizen Delibration.” in *Cambridge Handbook of Experimental Political Science*, J. N. Druckman, D. P. Green, J. H. Kuklinski and A. Lupia (eds.), Cambridge University Press, 2011, pp. 258-272.

Lazear, Edward P., and Kathryn L. Shaw. 2007. “Personnel Economics: The Economist’s View of Human Resources.” *Journal of Economic Perspectives* 21(4): 91-114.

Le Bon, Gustave, 1895/2009. *The Psychology of the Crowd*. Southampton: Sparkling Books.

Levine, David, and Thomas Palfrey. 2007. “The Paradox of Voter Participation? A Laboratory Study.” *American Political Science Review* 101 (01): 143–58.

Long, Richard J. 1978. “The effects of employee ownership on organizational identification, employee job attitudes, and organizational performance.” *Human Relations*, 31(1), 29–48.

Lopez de Leon, Fernanda, and Renata Rizzi. 2014. “A Test for the Rational Ignorance Hypothesis: Evidence from a Natural Experiment in Brazil.” *American Economic Journal: Economic Policy* 6 (4): 380-398.

Mazar, Nina and Chen-Bo Zhong (2010), “Do Green Products Make Us Better People?” *Psychological Science* 21(4): 494-498.

Mellizo, Philip, Jeffrey Carpenter, and Peter Hans Matthews. 2014. “Workplace democracy in the lab.” *Industrial Relations Journal* 45 (4): 313–328.

Messer, Kent, Homa Zarghamee, Harry Kaiser and William Schulze. 2007. “New hope for the voluntary contribution mechanism: the effects of context”, *Journal of Public Economics* 91, 1783-1799.

Mill, John Stuart. 1861. *Considerations on Representative Democracy.* iBook.

Miller, David. 1992. “Deliberative Democracy and Social Choice.” *Political Studies* 40 (s1): 54-67.

Mukand, Sharun and Rodrik, Dani. 2015. “The Political Economy of Liberal Democracy”, WP July.

Mullen, Elizabeth and Monin, Benoît. 2016. “Consistency Versus Licensing Effects of Past Moral Behavior,” *Annual Review of Psychology*, 67 (1), 363-85.

Page, Scott E. 2007. *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*. Princeton: Princeton University Press.

Pateman, Carole. 1970. *Participation and Democratic Theory*. Cambridge: Cambridge University Press.

Przeworski, Adam, and Fernando Limongi. 1993. "Political Regimes and Economic Growth." *Journal of Economic Perspectives* 7 (3): 51-70.

Simon, Adam F., and Tracy Sulkin. 2002. "Discussion's Impact on Political Allocations: An Experimental Approach." *Political Analysis*, 10 (4): 402-411.

Sulkin, Tracy, and Adam F. Simon. 2001. "Habermas in the Lab: A Study of Deliberation in an Experimental Setting." *Political Psychology*, 22 (4), 809-826.

Surowiecki, James. 2004. *The Wisdom of Crowds*. New York: Anchor Books.

Sutter, Mathias, and Christina Strassmair. 2009. "Communication, cooperation and collusion in team tournaments—An experimental study." *Games and Economic Behavior* 66(1): 506-525

Sutter, Mathias. 2009. "Individual behavior and group membership: comment." *American Economic Review* 99(5): 2247-2257.

Sutter, Matthias, Stefan Haigner, and Martin G. Kocher. 2010. "Choosing the carrot or the stick? Endogenous institutional choice in social dilemma situations." *Review of Economic Studies* 77(4): 1540-1566.

Van den Steen, Eric. 2009. "Authority versus Persuasion." *American Economic Review* 99 (2): 448–453.

Walker, John, Roy Gardner, Andrew Herr and Elinor Ostrom. 2000. "Collective choice in the commons: experimental results on proposed allocation rules and votes." *Economic Journal* 110, 212-234.